

### 多眼カメラを用いた超解像技術

HIRAO, Daiki / 平尾, 大樹

---

(出版者 / Publisher)

法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編 / 法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

57

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2016-03-24

(URL)

<https://doi.org/10.15002/00013286>

# 多眼カメラを用いた超解像技術

## PROTOTYPE OF SUPER-RESOLUTION CAMERA ARRAY SYSTEM

平尾 大樹

Daiki HIRAO

指導教員 彌富 仁

法政大学大学院理工学研究科応用情報工学専攻修士課程

We present a prototype of a super-resolution camera array system. Since the proposed system consists of 12 low-cost camera devices, all of which operate synchronously. It is a low-cost and high quality imaging system. However, when the targets are located near the system, parallax and differences in photographic conditions among the cameras become pronounced. In addition, conventional super-resolution techniques frequently emphasize noise, when the number of the observed images is limited. Therefore, we propose three ideas for our camera-array system. We confirm that the proposed system in general reduces the drawbacks of the array system and achieves approximately a 2 dB higher S/N ratio.

**Keywords** : Super-resolution, camera, image processing

### 1 はじめに

超解像技術とは観測した低解像度画像を用いて高解像度画像を生成する技術を指し、近年数多くの研究が報告されている [1]。この超解像技術は大きく2つに分けることができる。1つは学習型と呼ばれる超解像技術である。この手法は1枚の観測画像だけを用いて、補間法では再現できないエッジ等の細部を予め学習したデータベースを参照することで推定・復元する方法 [2] や、補間の考え方に基づく信号処理的なアプローチ [3] が提案されている。この方法を使えば学習に基づき画像の高周波成分を付加できる可能性があるが、学習した結果から参照しているため、必ずしも真の高周波成分を得られるとは限らないという欠点が挙げられる。

もう1つは再構成型と呼ばれる超解像技術である。この手法は複数枚の入力画像を用いて、処理後の画像の画素密度を高めていくことで解像度を向上させる方法である。この代表的な手法として、ML(Maximum likelihood)法、MAP(Maximum a posteriori)法 [4][5]、IBP(Iterative backward projection)法 [6]、周波数領域手法 [7]、POCS(Projection onto convex sets)法 [8][9]などがあげられる。学習型とは違い、この再構成型は真の高周波成分を得ることができる。本研究では最終目的として多眼カメラによる監視カメラの制作を目指しているため、観測画像内の正しい情報を得ることが重要である。そこで、真の高周波成分を得ることができる再構成型の超解像処理を扱い、その中でも最尤推定の原理に基づくML法を使用する。しかし、この方法は複数枚の観測画像を使用するため、再構成処理の前処理として、観測の

際の画像間の位置ずれを直す必要がある。この時、画像間の位置ずれは平行移動のみとは限らず、回転や拡大縮小などの複雑な変形の場合がある。位置合わせ処理には、画像内の特徴点を抽出してその対応点を利用する方法や、領域ベースマッチングなどの様々な手法 [10] がある。再構成処理による高画質化を実現するためには、サブピクセル精度の精密な位置合わせ処理が必要となり、超解像処理における課題の1つと言える。

一般的に超解像処理は1つのカメラから撮影された複数枚の観測画像を使用するため、撮影対象に動いている物体が存在していた場合、複雑な位置合わせ処理が必要となる。そこで、本研究では複数の低コストな撮像デバイスをアレイ状に配置し、同期的に制御するシステムを提案する。これにより高額な光学システムを導入することなく高解像画像が得られるほか、特別な前処理などを必要としないで動体に対する高解像化も期待できる。しかし、多眼カメラによる同時撮影を利用する場合 (1) カメラごとの色合いの違い、(2) カメラごとの視差、(3) 撮影枚数がカメラの台数に依存するなどの欠点が挙げられる。そこで本研究では、これらの問題に対する対策も合わせて提案する。また、監視カメラでは様々な場面に対応する必要があるため、撮影対象までの距離が遠い場合、撮影対象までの距離が近い場合、撮影環境が暗い場合を想定し、実際に多眼カメラを用いて撮影をして処理を行う。

### 2 超解像処理

本研究では複数枚の観測画像を用いた再構成型超解像処理について扱う。超解像処理は主に下記の2つの処理によって成り立っている。

1. 位置合わせ処理
2. 再構成処理

以下詳細について記述する。

(1) 位置合わせ処理

再構成型の超解像処理ではサブピクセル精度の正確な位置合わせ処理が要求される。本研究における位置合わせ処理の流れを Fig.1 に示す。本研究では

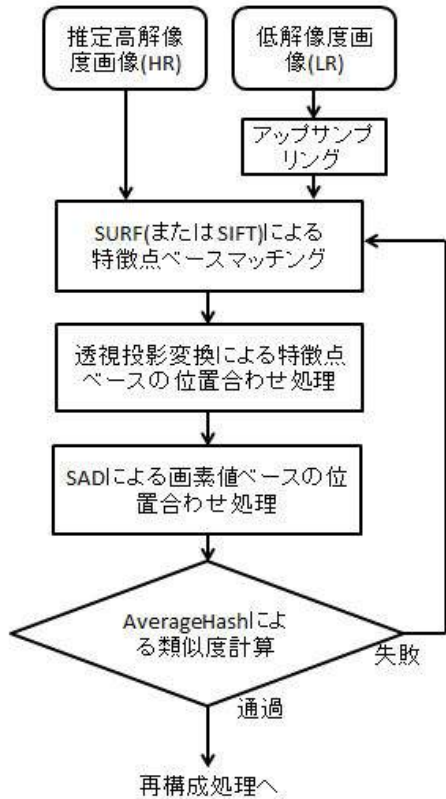


Fig.1 位置合わせ処理の流れ

大まかな位置合わせ処理に SURF(Speeded Up Robust Features)[11]、または SIFT(Scale-Invariant Feature Transform)[12] による特徴点マッチングを利用した。これらの手法は照明変化や回転、拡大縮小に不変な特徴量を取得する。しかし、SIFT は SURF に比べて特徴点検出の精度が良いが処理に時間がかかるという欠点がある。再構成型超解像処理は撮影した大量の観測画像すべてに対して位置合わせ処理を行うため、精度を重視して SIFT のみを使うと処理に多くの時間がかかってしまう。そのため、本研究では SURF と SIFT を組み合わせることにより、高精度かつ高速な位置合わせ処理を実現する。また、検出された画像間の対応点には一般的に不適切なものが含まれているため、これらを削除する必要がある。多眼カメラにより得られた画像群で同じ物体は似た位置に存在するという仮定のもと、ユーグリッド距離を用いて不適切な特徴点を検出して削除する。本研究では予備実験の結果、変動量の大きい特徴点の上位 5 割を削除して使用する。次に、残った信頼性の高い画

像間の対応点から、結んだ面積がより大きく、またより画像全体をカバーできるように 4 点を選択する。具体的には、残っている変動率の小さい特徴点の中で、画像の左上、左下、右上、右下の 4 隅からの距離が近い特徴点をそれぞれ選択して、透視投影変換を行う。

また、透視投影変換による位置合わせ処理を行った後、SAD(Sum of Absolute Difference) という輝度値の差の絶対値を算出する手法を用いて、画像全体を上下左右にピクセル精度でずらした画像を作成し、比較したい画像との絶対値の差が低い場所を捜し、画素値ベースの位置合わせ処理を行う。この 2 つの位置合わせ処理を組み合わせることにより、サブピクセル精度の精密な位置合わせ処理が実現できる。

これらの位置合わせ処理を行った後、画像間の類似度を評価し、類似度が低いと判断された場合、透視投影変換に用いる SURF 特徴点を変更して、再度位置合わせ処理を行う。また、繰り返し位置合わせ処理を行う回数は 4 回までとし、4 回目は SURF を用いていた特徴点検出を SIFT に切り替えて位置合わせ処理を行う。4 回とも位置合わせ処理を失敗してしまった場合、その画像は推定高解像度画像を更新するまでは使わないものとする。本研究では画像間の類似度計算法として Average Hash[13] を参考にした方法を用いた。Average Hash は、グレースケール化した画像を  $8 \times 8$  ピクセルに縮小化し、ピクセルの平均値を閾値として 2 値化することにより、 $8 \times 8 = 64$  ビットのビット列を生成する。そして、ビット列同士を比較することにより、画像間の類似度を高速に計算することができる手法である。本研究では予備実験の結果から、類似度計算の精度を高めるために、縮小化する正方形部位を  $30 \times 30$  ピクセルに変更して類似度を求めた。

(2) 再構成処理

本研究では、Farsiu らの手法 [14] を参考にして複数の低解像画像から高解像画像を得る再構成処理を行った。はじめに、画像の劣化モデルについて考える。元々は高解像度だった撮影対象が、画像として保存される段階において劣化してしまうのは、その過程で様々な影響を受けてしまうからだと考えられる。劣化の原因としてダウンサンプリング、光学系や大気による影響、カメラや被写体の移動による画素の移動などが挙げられる。この画像劣化モデルは (1) のように表現できる。

$$Y_k = D_k H_k F_k X, \quad k = 1, \dots, N. \quad (1)$$

$Y_k$  は観測低解像度画像、 $D_k$  はダウンサンプリング、 $H_k$  はぼけ関数、 $F_k$  は動きによる影響、 $X$  は推定高解像度画像、 $N$  は入力低解像度画像の枚数を表している。

不明な高解像度画像を、観測した低解像度画像、雑音やモーションのモデルに基づき ML 法を用いて推定す

る。ML 法は最尤推定の原理に基づく再構成型の超解像処理で広く使われる手法で、観測された低解像度画像から未知である推定高解像度画像が得られる確率が最大になるような推定高解像度画像を探索する方法である。この最小化問題を解決するために、最急降下法に基づいて (1) を書き直すと下記ようになる。

$$X_{n+1} = X_n - \left\{ \sum_{k=1}^N F_k^T H_k^T D_k^T \text{sign}(D_k H_k F_k X_n - Y_k) \right\}. \quad (2)$$

ここで、 $X_n$  は再急降下法によって生成された  $n$  番目の高解像度画像を示す。

### 3 多眼超解像処理システム

提案する多眼カメラのプロトタイプを Fig.2 に示す。このプロトタイプは 12 個のカメラ (AXIS M1011-W) から構成されていて、各カメラの解像度は  $640 \times 480$  ピクセルである。また、各カメラの視野角は水平方向が  $47^\circ$ 、垂直方向が  $35^\circ$  であり、プロトタイプ全体の大きさは  $18\text{cm} \times 24\text{cm}$  である。



Fig.2 多眼カメラ

多眼カメラは各カメラがアレイ状に配置されているため、観測画像ごとに視差の違いやピントのずれによる色合いの違いが起こる。そのため、本研究では多眼カメラによって撮影された画像を用いて超解像処理を行うために、2つの前処理をおよび画質改善策の計3つのアイデアを提案する。

#### (1) 色補正処理

本研究では任意に選択した推定高解像度画像を基準として、切り取った画像に対してヒストグラムマッチングによる色補正処理を行う。これにより、観測画像ごとの色合いを合わせてから超解像処理を行うことで、より高い解像度が望めると考えられる。

#### (2) 視差への対応

多眼カメラではカメラ間で視差が発生し、これは対象までの距離が近い場合に顕著に表れる。視差がある画像同士は、物体の奥行などの影響をうけ、正確な位

置合わせ処理は困難となるため、精密な位置合わせ処理が要求される超解像処理では視差の処理が重要になる。そのため、本提案システムでは解像度を高くしたい部分 (ROI) だけを半自動的に切り取って処理を行う方法を提案する。まず初めに、任意に選択した  $i$  番目の観測画像  $(1, 2, \dots, i, \dots, N)$  から  $ROI_i$  を手動で切り取る。この  $ROI_i$  を基に、2章で説明した特徴点マッチングを使用して画像間に対応する領域  $ROI_j (1, 2, \dots, j, \dots, N, j \neq i)$  を自動的に抽出する。カメラによって撮影された観測画像を Fig.3 に、切り取った観測画像を Fig.4 に示す。

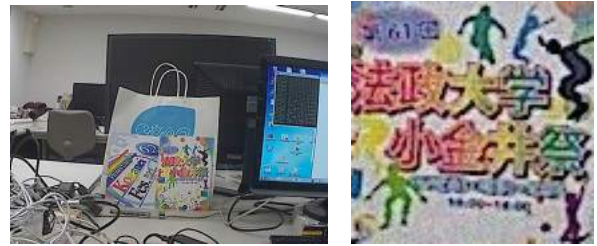


Fig.3 撮影した観測画像 Fig.4 切り取った観測画像

#### (3) 推定高解像度画像の雑音低減

通常、低コストカメラは小さいセンサー ( $1/4$ ,  $1/2.33$  など) を使用している。これらのセンサーを使用する場合、照明が不十分な環境で撮影を行うと、観測画像には雑音が付加されてしまうことが多い。また、多眼カメラ超解像システムは、観測画像の枚数がカメラの台数に依存するという欠点がある。使用する観測画像が少ない状態で超解像処理を行うと、画像内のエッジなどと同時に雑音も強調してしまう。そのため、以下の2つの処理を超解像プロセスに組み込むことを提案する。

本研究の超解像処理では、観測画像の1枚をアップサンプリングしたものを初期の推定高解像度画像として利用している。この時、まだ再構成処理を行っていない推定高解像度画像は、他の観測画像と同様に雑音が付加されている。そのため、従来の再構成型の超解像プロセスの初回に Median Filter を用いて推定高解像度画像の雑音除去を行うことを提案する。また、超解像処理に使用する観測画像が少ない場合、ML 法による推定高解像度画像の更新処理と同時に雑音が付加されてしまう可能性がある。そのため、本研究では推定高解像度画像と観測画像との誤差に対して Bilateral Filter を使用することによって、エッジを保持しつつ雑音を除去してから更新処理を行う。これらの方法を用いることによって、利用する低解像度画像の枚数が少ない場合でも、画像内の輪郭をぼやかすことをせずに雑音を除去することを期待できる。

## 4 結果

多眼カメラによって撮影した 12 枚の観測画像を用いて、作成した超解像処理を確認した。この時、処理を行った倍率は  $3 \times 3$  倍である。

まず初めに、遠距離を撮影した観測画像を用いて処理を行い、その結果を Fig.5 から Fig.7 に示す。それぞれ bi-linear を用いた結果 (Fig.5) と、12 枚の観測画像を用いた従来の超解像手法の結果 (Fig.6)、および提案手法の結果 (Fig.7) である。また、それぞれの一部拡大画像を Fig.8 から Fig.10 に示した。



Fig.5 遠距離撮影画像を用いた bi-linear 法



Fig.6 遠距離撮影画像を用いた従来超解像手法 (12 枚)



Fig.7 遠距離撮影画像を用いた提案超解像手法 (12 枚)

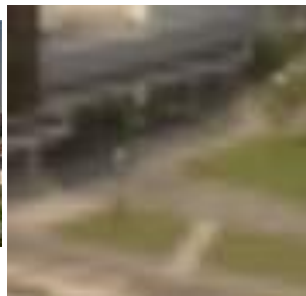


Fig.8 遠距離撮影画像を用いた拡大 bi-linear 法



Fig.9 遠距離撮影画像を用いた拡大従来手法 (12 枚)



Fig.10 遠距離撮影画像を用いた拡大提案手法 (12 枚)

次に、暗所を撮影した観測画像を用いて処理を行う。この時、撮影した暗所の観測画像に対し、ヒストグラムイコライザを行ってから超解像処理を行う。超解像処理の結果を Fig.11 から Fig.13 に示す。それぞれ bi-linear を用いた結果 (Fig.11) と、4 枚の観測画像を用いた従来の超解像手法の結果 (Fig.12)、および提案手法の結果

(Fig.13) である。また、それぞれの一部拡大画像を Fig.14 から Fig.16 に示す。



Fig.11 暗所撮影画像を用いた bi-linear 法



Fig.12 暗所撮影画像を用いた従来超解像手法 (4 枚)



Fig.13 暗所撮影画像を用いた提案超解像手法 (4 枚)



Fig.14 暗所撮影画像を用いた拡大 bi-linear 法



Fig.15 暗所撮影画像を用いた拡大従来手法 (4 枚)



Fig.16 暗所撮影画像を用いた拡大提案手法 (4 枚)

最後に、近距離を撮影した Fig.4 に対して処理を行った結果について、bi-linear を用いた結果と、12 枚の観測画像を用いた提案する超解像処理結果を Figs.17.18 に、同様に、少ない低解像度画像しか得られない場合の比較のために、4 枚の画像を用いて従来手法並びに提案手法で再構成した結果 Figs.19.20 に、また、より分かりやすくするために、これらの一部拡大画像を Fig.21 から Fig.24 に示す。

従来手法と提案手法の数的評価のために、40 枚の観測画像を利用した超解像処理の結果を真と仮定した時、使用した観測画像の枚数と超解像手法による SN 比の変化を Fig.25 に示す。

以上の結果から、従来手法に比べて提案手法の結果の方が雑音が少ないことがわかる。



Fig.17 近距離撮影画像を用いた bi-linear 法



Fig.18 近距離撮影画像を用いた提案超解像手法 (12枚)



Fig.19 近距離撮影画像を用いた従来手法 (4枚)



Fig.20 近距離撮影画像を用いた提案手法 (4枚)



Fig.21 近距離撮影画像を用いた拡大 bi-linear 法



Fig.22 近距離撮影画像を用いた拡大提案手法 (12枚)

## 5 考察

まず初めに、遠距離を撮影した画像を用いて処理を行った場合を考える。バイリニア法を用いて拡大した結果と比べてみると、多眼カメラを用いて撮影した画像に対して超解像処理が行われていることがわかる (Figs.5.6)。しかし、従来の手法を用いた結果と、提案する手法を用いた結果の視覚的に優位な差は確認できなかった (Figs.6.7)。これは、観測画像内のノイズが少なく、再構成に使用した画像の枚数も多かったためだと考えられる。このことは、それぞれの結果を拡大した画像 (Figs.8~.10) からよくわかる。



Fig.23 近距離撮影画像を用いた拡大従来手法 (4枚)



Fig.24 近距離撮影画像を用いた拡大提案手法 (4枚)

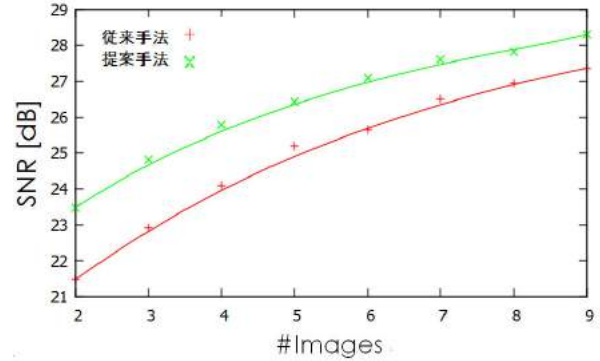


Fig.25 近距離撮影画像を用いた超解像処理結果の S/N 比

次に、暗所を撮影した画像を用いて処理を行った場合を考える。バイリニア法を用いて拡大した結果と比べてみると、超解像処理結果の方が文字が読み取りやすくなっていることがわかる (Figs.11.12)。しかし、超解像処理結果は同時にノイズも強調してしまっている。これは、超解像処理に使用した観測画像の枚数が4枚と少なく、観測画像に多くのノイズが含まれているためであると考えられる。このことは、それぞれの結果を拡大した画像 (Figs.14.15) からよくわかる。また、従来手法 (Fig.12) と提案手法 (Fig.13) を比較してみると、どちらの手法も同じぐらい文字が強調されているが、従来手法に比べ、提案手法の方がノイズが少ないことがわかる。このことから、ノイズ低減を目的とした2つのフィルターを新たに加えた提案手法が上手く機能していると考えられる。

最後に、近距離を撮影した観測画像を用いて処理を行った場合を考える。この場合、近距離撮影の欠点である視差や色合いの問題を考慮することで、こちらも問題なく解像度を向上することができた (Figs.21.22)。

次に、観測画像が少ない場合を比較してみると、従来手法 (Fig.23) ではエッジが強調されていると同時に、雑音も強調されてしまっていることがわかる。それに対し、提案手法 (Fig.24) では従来手法 (Fig.23) と同様にエッジが強調されているが、雑音は少ない状態であることがわかる。このことから、本研究で提案する手法は従

来の超解像手法と同等以上の高解像度化が望め、雑音には強い手法であることがわかる。

また、S/N 比の比較 (Fig.25) から、提案手法は従来手法に用いた観測画像の枚数プラス 2 枚分程度の SN 比が得られている。

以上の結果から、提案する手法は雑音に強く、撮影枚数が少ない場合でも良い超解像処理の結果が得られるのではないかと考えられる。

## 6 まとめ

今回、本研究では多眼カメラのプロトタイプを作成し、撮影をして超解像処理を行った。提案する多眼カメラを用いた超解像処理として、(1) 観測画像に対しての色補正処理、(2) 視差が大きい場合の ROI 処理、(3) 観測画像の枚数が少ない場合のフィルター処理の 3 つの改善を取り入れることにより、従来手法よりも 2dB 程 (観測画像約 2 枚分程度) の良い結果を得ることができた。今後は多眼カメラの小型化と共に、さらなる高解像度化を目指していきたい。

## 7 謝辞

本研究にあたり、全般にわたるご指導並びにデータセットの提供をしてくださった彌富仁准教授、および彌富研究室の皆様に深く御礼申し上げます。

## 参考文献

- 1) Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: A technical overview. *IEEE Signal Process. Mag* 20 (2003) 21-36
- 2) Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-based super-resolution. *IEEE Computer Graphics and Applications* 22 (2002) 56-65
- 3) Takashi, K., Takahiro, S.: Super-resolution decoding of the jpeg coded image data using total-variation regularization. *Picture Coding Symposium* (2010) 114-117
- 4) Schultz, R., Stevenson, R.: A bayesian approach to image expansion for improved definition. *IEEE Transactions on Image Processing* 3 (1994) 233-242
- 5) Schultz, R., Stevenson, R.: Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing* 5 (1996) 996-1011
- 6) Zitova, B., Flusser, J.: Image registration methods: A survey. *Image and Vision Computing* 21 (2003) 977-1000
- 7) Tsai, R., Huang, T.: Multiframe image restoration and registration. *Advances in Computer Vision and Image* 1 (1984) 317-339
- 8) Eren, P., Sezan, M., Tekalp, A.: Robust, object-based high resolution image reconstruction from low-resolution video. *IEEE Transactions on Image Processing* 6 (1997) 1446-1451
- 9) Patti, A.J., Altunbasak, Y.: Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *IEEE Transactions on Image Processing* 10 (2001) 179-186
- 10) Zitova, B., Flusser, J.: Image registration methods: A survey. *Image and Vision Computing* 21 (2003) 977-1000
- 11) Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (surf). *Computer Vision and Image Understanding* 110 (2008) 346-359
- 12) Lowe, D.: Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision* (2004) 91-110
- 13) Aghav, S., Kumar, A., Gadakar, G., Mehta, A., Mhaisane, A.: Mitigation of rotational constraints in image based plagiarism detection using perceptual hash. *International Journal of Computer Science Trends and Technology* 2 (2014) 28-32
- 14) Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multi-frame super-resolution. *IEEE Transaction on image processing* 13 (2004) 1327-1344